Federal Register Notice 87 FR 35250, https://www.federalregister.gov/documents/2022/06/09/2022-12432/request-for-information-on-advancing-privacy-enhancing-technologies, June 9, 2022

Request for Information (RFI) on Advancing Privacy Enhancing Technologies

Amazon Web Services (AWS)

DISCLAIMER: Please note that the RFI public responses received and posted do not represent the views or opinions of the U.S. Government. We bear no responsibility for the accuracy, legality, or content of the responses and external links included in this document.

July 8, 2022

Re: Request for Information on Privacy Enhancing Technologies

Submitted To:	**Submitted By:**
The White House Office of Science and Technology Policy *Via Email*	Amazon Web Services, Inc. 12900 Worldgate Dr. Herndon, VA 20170

Amazon Web Services (AWS) appreciates the opportunity to submit feedback to the Office of Science and Technology Policy (OSTP) in response to its Request for Information (RFI) on Advancing Privacy-Enhancing Technologies (PETs).[1]

This RFI seeks input on technologies that advance mechanisms for privacy-preserving data sharing and analytics technologies to inform a national strategy that will put forth a vision for responsibly harnessing such technologies to benefit individuals and society. Some of the technologies and techniques referenced in the RFI include differential privacy, homomorphic encryption, and secure enclaves. AWS has invested considerable efforts into many of these technologies and techniques. We are grateful for this opportunity to provide feedback to OSTP on our efforts to help inform the national strategy that will help create an environment in which such technologies can be more broadly adopted and flourish to harness the power of data.

AWS is the world's most comprehensive and broadly adopted cloud platform, and serves millions of customers who are primarily businesses, non-profits, and government organizations. In this response we focus on efforts that AWS has undertaken to enable our customers with privacy-enhancing technologies. In addition to directly investing in PETs, AWS is also focused on related privacy topics including data governance, lineage, access controls, audits, and retention policies, among other areas. As a result of evolving privacy expectations of individuals, companies, and governments, the privacy needs of our customers are expanding rapidly, including the need for technologies that allow for disassociation and confidentiality. The way each company implements data privacy is dependent on their business goals; our goal is to provide customers a comprehensive set of technical capabilities that they can use to implement privacy into their architectures depending on their business needs and risk posture.

In our comments below, we describe specific PETs that we believe OSTP should focus on in its national strategy, as well as specific applications and research efforts that could help inform the national

[1] https://www.federalregister.gov/documents/2022/06/09/2022-12432/request-for-information-on-advancing-privacy-enhancing-technologies

strategy, specific applications of these PETs in AWS products and services, as well as noting relevant limitations, tradeoffs, and barriers to adoption. Finally, we have set forth regulatory efforts that we believe can help in ensuring a thriving market for the creation, adoption, and use of PETs in the future.

Specific Privacy-Enhancing Technologies that Advance Greater Data Analysis and Sharing

Below we describe some PETs that AWS has focused on that can be leveraged to de-identify and disassociate data for the purpose of data analysis and sharing, and which we believe OSTP should also focus on in its national strategy. Increasingly, merging and leveraging different data sets can facilitate innovation and benefit the public. We have provided examples in each area of how such technologies have been utilized or could be utilized to enable data collaboration amongst multiple parties while disassociating the data from the data subject and preserving privacy.

Differential Privacy

We believe differential privacy (DP) is a powerful technique for protecting individuals' privacy when their data is used to derive insights about a data set, and OSTP should closely examine and invest further in this technology. Amazon Science has published a number of articles on differential privacy, which can be found here: https://www.amazon.science/tag/differential-privacy

The main idea behind DP is to add randomness — noise— to obscure the presence or absence of any single individual in the data set. Randomness can be integrated at different stages of an algorithm. For example, noise can be added to the data itself or to the output of the queries computed on the data. The more noise that is introduced, the greater the privacy protection, but more noise also degrades the utility of the algorithm outputs. This trade-off between protecting the privacy of users and providing useful insights about the population is controlled by a parameter known as *privacy budget*. The privacy budget is a finite resource that is consumed each time a query is run. Differential privacy is best suited for use cases where a small amount of error in the results of a query is acceptable, and does not destroy the utility of the query results. At AWS, we are actively conducting research and development into differential privacy technologies that make secure computation easier and more efficient. We believe this technique is an important building blocks for better data governance and privacy, and that there is significant privacy value to be gained from investments in this area.

Ongoing research on privacy-preserving machine learning in natural language models is summarized in this Amazon Science blog post: https://www.amazon.science/blog/advances-in-trustworthy-machine-learning-at-alexa-ai. As noted in this blog post, differential privacy provides a rigorous way to quantify the privacy of machine learning models. The researchers investigated vulnerabilities presented in the differential-privacy literature and proposed computationally efficient mechanisms for protecting against them. Additionally, training machine learning models on synthetic, differentially private data can be a powerful way to make machine learning more privacy protective; these efforts are detailed further in this blog post: https://www.amazon.science/blog/amazon-helps-launch-workshop-on-synthetic-data-generation.

Differential privacy technology can be applied to enable multiple parties to engage in data collaborations. In data collaborations, one party (data consumer) can join their data with that of another party (data provider), and run aggregate queries (e.g. sum, count, average) against the joint dataset. These collaborations are commonly known as 'clean rooms' because neither party's raw data is revealed to each other in the process. AWS customers, particularly in the online advertising context, can drive privacy protective multi-party data collaboration using data clean rooms: https://aws.amazon.com/blogs/industries/deploying-a-privacy-safe-data-clean-room-on-aws-with-snowflake/.

Secure Enclaves

AWS Nitro Enclaves enables customers to create isolated compute environments to further protect and securely process highly sensitive data such as personally identifiable information (PII), healthcare, financial, and intellectual property data within their Amazon EC2 instances.

Nitro Enclaves helps customers reduce the attack surface area for their most sensitive data processing applications. Enclaves offers an isolated, hardened, and highly constrained environment to host security-critical applications. Nitro Enclaves enables a range of use cases that deal with the processing of highly sensitive data, such as securing private keys, tokenization, and multi-party collaboration. Nitro Enclaves' isolation, cryptographic attestation capabilities, and integration with AWS Key Management Service, are key features that provides customers with a practical approach to setting up multi-party collaboration. To illustrate how multiparty collaboration can work with Nitro Enclaves, we have recently published an example proof of concept (POC) on third-party bidding service for real estate transactions. In the POC, buyers will submit encrypted bids to the application. Once all the bids have been entered, the application will decrypt the bids, determine the highest bidder, and return a result without disclosing the actual bid amounts to any party. For more details, see: https://aws.amazon.com/blogs/compute/leveraging-aws-nitro-enclaves-to-perform-computation-of-multiple-sensitive-datasets/.

Computing Over Encrypted Data

AWS Cryptography tools and services utilize a wide range of encryption and storage technologies that can help customers protect their data at rest and in transit. In some instances, customers also require protection of their data even while it is in use. To address this need, AWS is developing new techniques for cryptographic computing, an emerging technology that allows computations to be performed on encrypted data, so that sensitive data is never exposed. It can be the foundation used to help protect the privacy and intellectual property of data owners, data users, and other parties involved in machine learning activities. AWS presented on this class of techniques at our re:Invent conference in 2020, which can be accessed here, along with other cryptographic computing publications: https://aws.amazon.com/security/cryptographic-computing/.

However, the computational resources and cost needed for this type of computing pose significant barriers to adoption. We believe this is an area of research that could benefit greatly from OSTP investment to materialize and operationalize.

Other cryptography efforts are also worth noting. Specifically, AWS has focused significant efforts on post-quantum cryptography for the cloud, and some of our efforts in this area can be found here: https://d1.awsstatic.com/events/reinvent/2020/Building_PostQuantum_Cryptography_for_the_Cloud_SEC207.pdf. In 2020, AWS completed benchmarks of Round 2 Versions of the Bit Flipping Key Encapsulation (BIKE) and Supersingular Isogeny Key Encapsulation (SIKE) hybrid post-quantum Transport Layer Security (TLS) Algorithms, which were submitted to the National Institute of Standards and Technology (NIST) as part of NIST's Post-Quantum Cryptography standardization process: https://aws.amazon.com/blogs/security/round-2-hybrid-post-quantum-tls-benchmarks/

Federated Learning

Federated learning allows data from multiple sources to be examined while limiting researcher access to each data source to help preserve privacy. For example, in some federated approaches, data contained in disparate sources can be queried for information on a certain question and only aggregated or de-identified data is returned to the researcher. Federation also enables the development of scalable architectures, where new data sources can be added to research models. As a result, use of federated systems can support initial proof-of-concepts that can then grow to include new data sources and further accelerate the time-to-science. Here is an example of the use of federated analysis in the European Healthcare system: https://aws.amazon.com/blogs/industries/data-mesh-technology-can-enable-european-healthcare-system-collaborate/

Regulatory Recommendations

PET adoption may be enhanced through meaningful measurement standards. We believe a risk-based approach which accounts for different ways to mitigate risks to acceptable levels, and different ways to measure that risk, would be a beneficial approach. It is important to address measurement standards from a flexible, risk-based approach (e.g. different models will reduce risk to different levels) to account for the vast number and type of contexts in which PETs would be deployed. Furthermore, data that has been de-identified should not be within the scope of privacy laws or regulations as disassociation would remove potential risks to individuals related to the data. This will also continue to promote the faster adoption of PETs as a way for organizations to reduce their privacy risks, and also retain the flexibility for the development of new and enhanced technologies.

We encourage OSTP to support efforts for sensible, comprehensive data privacy protections in the U.S. Such efforts should not deter innovation, and should encourage the development and adoption of techniques and controls that obstruct reidentification. Any such legislation should also clearly understand the different roles of data controllers and processor. Notably, efforts to ensure responsible and fair use of artificial intelligence, which is an application for many of the technologies above, should remain in a separate framework.

In addition, we encourage OSTP to work with each federal agency to accelerate adoption of PETs—particularly for those agencies that process more sensitive data. For example, OSTP should work with health agencies to accelerate use of PETs to spur biomedical innovation and provide patients with greater ability to direct the sharing of their data. OSTP should also work with the Department of Health and Human Services (HHS) to accelerate individual-level control and input on the sharing of their data. In this way, individuals could authorize the use of their data for different purposes, such as sharing their data with other care providers, discovering clinical trials that may be relevant for them, or volunteering data for research purposes. Cross-agency collaboration—including with the Centers for Medicare & Medicaid Services, Office of the National Coordinator for Health Information Technology, National Institutes of Health, and Office of Civil Rights—can identify opportunities to accelerate greater individual-level control. OSTP should work with HHS to examine the development of incentives (e.g., via electronic health record certification or the Promoting Interoperability program) and other policies to encourage healthcare providers to support individual-level control on the sharing of their data.

We thank you for the opportunity to respond and provide input into the development of the national strategy for advancing PETs.

Sincerely,

Shannon Kellogg

Vice President, AWS Public Policy --- Americas